Between and within-herd variation in blood and milk biomarkers in Holstein cows in early lactation

Between and within-herd variation in blood and milk biomarkers in Holstein cows in early lactation

M.A. Krogh[1], M. Hostens [2], M. Salavati [3,a], C. Grelet [4], M.T. Sorensen [1], D.C. Wathes [3], C.P. Ferris[5], C. Marchitelli [6], F. Signorelli[6], F. Napolitano[6], F. Becker [7], T. Larsen [1], E. Matthews [8], F. Carter [8], A. Vanlierde[4], G. Opsomer [2], N. Gengler [9], F. Dehareng[4], M.A. Crowe [8], K.L. Ingvartsen[1] and L. Foldager[1,10]

[1] *Department of Animal Science, Aarhus University, Blichers Allé 20, DK8830 Tjele, Denmark*

[2] *Department of Reproduction, Obstetrics and Herd Health, Faculty of Veterinary Medicine, Ghent University, 9820 Merelbeke, Belgium*

[3] *Royal Veterinary College, Hawkshead Lane, North Mymms, Hatfield, AL9 7TA, United Kingdom*

[4] *Valorization of Agricultural Products Department, Walloon Agricultural Research Center, 5030 Gembloux, Belgium*

[5] *Agri-Food and Biosciences Institute (AFBI), Belfast, Northern Ireland*

[6] *Research Center for Animal Production and Aquaculture (CREA), Roma, Italy*

[7] *Leibniz Institute for Farm Animal Biology (FBN), Dummerstorf, Germany*

[8] *University College Dublin (UCD), Dublin, Ireland*

[9] *Gembloux Agro-Bio Tech, University of Liège, Gembloux, Belgium*

[10] *Bioinformatics Research Centre, Aarhus University, C.F. Møllers Allé 8, DK8000 Aarhus, Denmark*

[a] *Present address:* Genetics and Genomics Division, The Roslin Institute

Easter Bush Campus, Midlothian, EH25 9RG, United Kingdom

Corresponding author: Mogens A. Krogh. Email: mogenskrogh@anis.au.dk

Short title: Variations in early lactation biomarkers

Abstract

Both blood- and milk-based biomarkers have been analysed for decades in research settings, although often only in one herd, and without focus on the variation in the biomarkers that are specifically related to herd or diet. Biomarkers can be used to detect physiological imbalance and disease risk, and may have a role in precision livestock farming (**PLF**). For use in PLF it is important to quantify normal variation in specific biomarkers and the source of this variation. The objective of this study was to estimate the between and within-herd variation in a number of blood metabolites (β-hydroxybutyrate (**BHB**), non-esterified fatty acids, glucose and serum IGF-1), milk metabolites (free glucose, glucose-6-phosphate, urea, isocitrate, BHB and uric acid), milk enzymes (lactate dehydrogenase and *N*-acetyl-β-D-glucosaminidase (**NAGase**)) and composite indicators for metabolic imbalances (Physiological Imbalance-index (**PI-index**) and energy balance (**EBAL**)), to help facilitate their adoption within PLF. Blood and milk were sampled from 234 Holstein dairy cows from six experimental herds, each in a different European country, and offered a total of 10 different diets. Blood was sampled on two occasions at approximately 14 days-in-milk (**DIM**) and 35 DIM. Milk samples were collected twice weekly (in total 2750 samples) from DIM 1 to 50. Multilevel random regression models were used to estimate the variance components and to calculate the Intra-Class Correlations (**ICC**). The ICCs for the

milk metabolites, when adjusted for parity and DIM at sampling, demonstrated that between 12% (glucose-6-phosphate) to 46% (urea) of the variation in the metabolites' levels could be associated with the herd-diet combination. ICC related to the herd-diet combination were generally higher for blood metabolites; from 17% (cholesterol) to approximately 46% (BHB and urea). The high ICC for for urea suggest that this biomarker can be used for monitoring on herd level. The low variance within cow for NAGase indicate that few samples would be needed to describe the status and potentially a general reference value could be used. The low ICC for most of the biomarkers and larger within cow variation emphasizes that multiple samples would be needed most likely on the individual cows for making the biomarkers useful for monitoring. The majority of biomarkers were influenced by parity and DIM which indicate that these should be accounted for if the biomarker should be used for monitoring.

Keywords: dairy, biomarker, physiological imbalance, variance, monitoring

Implications

We quantified normal variation in blood- and milk-based biomarkers of health and performance among cows housed in very different feeding and housing conditions. Some biomarkers like urea were strongly affected by herd factors, which make them very useful for herd-level monitoring. Some biomarkers like *N*-acetyl-β-D-glucosaminidase (NAGase) were very uniform from day to day at cow level, which make it possible to monitor health and performance with few samples and sometimes to use general reference values. Most biomarkers were substantively affected by

factors like parity and stage of lactation, which demonstrates that these factors must be accounted for in monitoring programs.

Introduction

Modern dairy farming faces many challenges, including the need to optimise production efficiency, while at the same time maintaining the physiological balance, health and fertility of the cows. High yielding dairy cows in early lactation frequently suffer from their inability to consume sufficient feed to support the amount of energy required for maintenance and milk production, leading to negative energy balance (**EBAL**), which is reflected in the mobilisation of body reserves. If this mobilisation is too extensive, this may create physiological imbalances, which in turn make the animal more susceptible to both metabolic and infectious diseases (Ingvartsen, 2006). To date, most indicators of physiological imbalance have been based on measurements of blood metabolites (e.g. β-hydroxybutyrate (**BHB**), non-esterified fatty acids (**NEFAs**), glucose and urea), and hormones such as serum IGF-1 (Chagas *et al.*, 2007; Wathes *et al.*, 2007; Ingvartsen and Moyes, 2013). Milk enzymes such as *N*-acetyl-β-D-glucosaminidase (**NAGase**) and lactate dehydrogenase (**LDH**) also show potential as biomarkers to describe udder health. Currently tests for BHB in milk and blood and blood glucose (Mair *et al.*, 2016) and LDH can be conducted cow-side.

New spectrophotometric techniques such as Fourier transform mid-IR (**FT-MIR**) spectra of milk show promising perspectives to provide more timely information for improved herd management in relation to metabolic health. FT-MIR can reliably predict the concentrations of BHB, acetone and citrate (Grelet *et al.*, 2016a). It is highly likely that other metabolites can also be predicted from FT-MIR, and this would

reduce the costs (and possibly time) associated with more traditional biochemical methods and provide the possibility to have more detailed time series of observations. Such developments lead to the potential for incorporating the monitoring of metabolic health into the area of precision livestock farming (**PLF**). Berckmans (2017) defined PLF as managing individual animals by continuous real-time monitoring of health, welfare and production. In non-biological production systems, process control charts are often used to monitor the production system, however they have also been applied in animal production (Mertens *et al.*, 2011). One approach of what should be monitored in process control is that it is not the outcome of the production process that is monitored but a suitable indicator of the production process that mirrors the current state of the production system (Mertens *et al.*, 2011). The goal is to have a production process where the amount or quality of the process outcome will be predictable using the indicator. Despite a production process being predictable, it does not necessarily mean that it is also acceptable. Using this terminology with respect to the present paper, process outcome refers to the occurrence of metabolic diseases and indicators refer to the observations of the individual cows' metabolic status measured by the biomarkers. The production process is then predictable if the biomarkers can predict the occurrence of metabolic diseases, but if not predictable, assignable cause of process variation should be eliminated. Process control charts were a natural predecessor of PLF although to develop both, detailed information about the normal variation in the control input is needed. Milk and blood biomarkers have been analysed for decades in research settings (Andersson, 1984; Jensen *et al.*, 1993) and in some commercial herds (Stengärde *et al.*, 2010). Very often the observations come from one research herd, making the estimation of herd variance impossible. When multiple herds are

126 included, herd effects are often treated as uninteresting covariates that are all too
127 often removed to elucidate the research question (e.g. Ospina *et al.*, 2010; Seifi *et*
128 *al.*, 2011). If substantial between- and within-herd variation exists in blood and milk
129 biomarkers, this may result in poor predictions of the process outcome. The four
130 different levels of variation (random, within-cow/temporal, cow-level and herd-level)
131 are described as follows. 1) Random variation, due to measurement error induced by
132 sampling or analytical error. 2) The temporal within cow variation in the biomarker:
133 this is due to diurnal variation (Nielsen *et al.*, 2003) and differences in sampling time
134 relative to feeding. Both random and within cow variation should be minor compared
135 to the difference in means between the healthy and imbalanced population,
136 otherwise there will be substantial misclassification of the truly healthy versus truly
137 imbalanced cows. This could be accounted for by more observations from each cow
138 or by adjusting for diurnal patterns. 3) Variation in the biomarker due to differences
139 between cows: this source of variation arises from cows responding differently to the
140 same metabolic challenge, even if they remain healthy. Genetic differences between
141 individual cows that result in some metabolic pathways functioning more efficiently
142 than others could result in systematic differences in baseline levels of the
143 biomarkers. Such genetic differences in EBAL have been found (Berry *et al.*, 2007).
144 Also, the cow's physiological constitution could provide systematic differences in the
145 biomarkers: for example two cows with the same challenge in terms of EBAL would
146 most likely respond very differently in blood/milk BHB depending on their body
147 condition. This is related to construct validity (O'Leary-Kelly and Vokurka, 1998)
148 where it should be carefully evaluated if the observed biomarker actually measure
149 what it is intended to –here the metabolic challenge the cow is exposed to. The
150 importance of cow-to-cow variation within healthy or imbalanced cows depends on

151 the size of the variation, compared to differences between balanced and imbalanced
152 cows. If cow-to-cow variations are large within healthy cows, then biomarker
153 observations that are normal for one cow could be imbalanced for another, thus
154 requiring multiple observations on individual cows for PLF tools to function. Additional
155 observations of the individual cow will not improve this issue, because repeated
156 observations within an individual cow will be correlated. 4) Variation in individual cow
157 biomarkers that can be attributed to all cows in the herd: for example, feeding
158 strategy will clearly affect the metabolic status of cows in any herd, and dependent
159 on the diet, a small or large number of cows may be imbalanced. Herd variation,
160 however, means that the same diet may result in specific patterns or levels of
161 biomarkers in different herds, dependent on other management factors such as
162 milking frequency, feed management, stocking density etc. Such management
163 factors could, for example, affect biomarkers by influencing the glucocorticoid levels
164 (Huzzey *et al.*, 2012) that influence energy metabolism. Another explanation for herd
165 variation could be the presence of dominant genotypes within herds, so it is actually
166 an accumulated cow-effect. The importance of the herd variation in biomarkers is
167 again dependent on differences in the biomarkers between the imbalanced and
168 healthy population. The implication of this sort of variation is that multiple
169 observations within each herd is needed to adjust the PLF-tool to the herd specific
170 threshold.
171 The objective of this study was to estimate the between- and within-herd
172 variation in key blood metabolites/hormones (BHB, NEFA, glucose, urea,
173 fructosamine, cholesterol and IGF-1), milk metabolites/enzymes (BHB, glucose, urea,
174 isocitrate, glucose-6-phosphate, uric acid, NAGase and LDH) and composite

indicators for metabolic imbalances (EBAL and PI-index), so as to assist the implementation of these into PLF.

Material and methods

Animals, data storage and transportation of samples

A comprehensive set of samples and data were collected between calving and 50 days post calving (1-50 days in milk (**DIM**)) from 241 Holstein cows in six research herds: 35 from **DK** (Aarhus University, Denmark); 39 from **IE** (UCD Lyons Research Farm, University College Dublin, Ireland); 62 from **UK** (Agri-Food and Biosciences Institute, Northern Ireland, UK); 31 from **BE** (Walloon Agricultural Research Centre, Belgium); 29 from **DE** (Leibniz Institute for Farm Animal Biology, Germany), and 45 from **IT** (Consiglio per la Ricerca in Agricoltura, Italy). Seven cows were culled before 50 DIM (one UK, three DE, and three IE) and were subsequently excluded, thus data from 234 cows were available for analyses. Of these 234 cows, 55 were in parity 1, 66 were in parity 2 and 113 were in parity ≥3 (**3+**), giving an overall median lactation number of 2 (max lactation number 7). All data were stored in a central repository at Dairy Data Warehouse, The Netherlands. Data were checked for errors, validity and agreement between original data and data from the repository. Blood plasma, serum and milk samples were transported frozen between the place of collection and the appropriate laboratory by commercial transport companies. Samples were shipped in insulated containers in the presence of dry ice. The consignments were in all instances checked for residual dry ice at reception.

Diets and feed intake

Cows in two of the herds (UK and DK) were offered three contrasting diets, some of which were designed to challenge the cows metabolically (the 'High sugar' diet should be ketogenic and the 'High starch' should induce acidosis in DK), while cows in the remaining four herds were offered diets which reflected local management practices (Table 1). Automated electronic feed intake recording systems were used to record daily intakes of individual cows from DK/IE (Insentec, Markneesse, Netherlands), and UK (Calan gates linked to an automatic cow identification system (American Calan, Northwood, NH), which allowed cows to gain access to feed boxes mounted on weighing scales (Griffith Elder, Bury St. Edmunds, UK)). In DE, the total mixed ration was placed in troughs on scales, connected to a computer. Feed intakes were not recorded in BE or IT.

Milk yield and composition

All cows were milked twice daily and yields (volumes) were recorded from approximately three DIM. Milk samples, containing Bronopol 0.02% as a preservative, were collected from consecutive morning and evening milkings twice weekly from seven DIM onwards, stored at 4°C and subsequently analysed by FT-MIR for composition of protein, fat and lactose. The morning and evening compositions were weighted for milk yields to provide a daily weighted average composition.

Live body weight recording and health records

Live body weights of cows were recorded on at least two occasions over the seven-week period from all herds except IT. The frequency of weighing differed markedly between herds: DK and UK at every milking (i.e. twice per day), IE and DE

on average two and three times per week, and BE twice during the study period (at approximately 14 and 35 DIM). Details of health problems and their treatments for individual cows were obtained from herd health records. The detection of health problems and subsequent treatment followed the normal management practice in the herds. No attempts were made to assess the agreement between diagnostic procedures in different countries.

Blood sampling and analyses of metabolites and IGF-1

At approximately 14 DIM (mean=14.1, SD=2.0, range: 11 to 20) and 35 DIM (mean=34.8, SD=1.9, range: 31 to 38), blood samples were collected by jugular venepuncture to obtain plasma (in Na heparin tubes) and serum (plain tubes): plasma and serum were separated by centrifugation, and stored at -20°C for subsequent analysis. Urea and cholesterol were determined in plasma according to standard procedures using an auto-analyser, ADVIA 1800 ® Chemistry System (Siemens Medical Solutions, Tarrytown, NY 10591, USA). Glucose, NEFA, BHB and fructosamine were determined according to Bjerre-Harpoth *et al.* (2016). *Intra*- and *inter*-assay CV were in all cases below three and four percent, respectively, for both low and high control samples. Laboratory analyses of all blood metabolites were carried out at the Department of Animal Science, Aarhus University, Denmark. Concentrations of IGF-1 were determined in serum by radioimmunoassay at University College Dublin, Ireland, following acid–ethanol extraction using the method as described by Beltman *et al.* (2010). *Intra*-assay CVs were 12.4, 7.5 and 9.9% for low, medium and high control samples, respectively. The corresponding *inter*-assay CVs were 7.8, 3.9 and 9.4%. The sensitivity of the assay, defined as the lowest concentration detectable, was 4 ng/ml.

Milk sampling and analysis for metabolites and enzymes

Additional milk samples were collected twice weekly during morning milking, starting at around seven DIM. On each occasion, two 8 ml samples were obtained and stored at -18°C in tubes with stoppers. Fluorometric end point analyses were used to determine milk glucose and glucose-6-phosphate (Larsen, 2015), uric acid (Larsen and Moyes, 2010), isocitrate (Larsen, 2014) and BHB (Larsen and Nielsen, 2005). Urea was determined by spectrophotometry (Nielsen *et al.*, 2005). The indigenous enzymes LDH (EC. 1.1.1.27) and NAGase (EC 3.2.1.30) were analysed by fluorometric assays according to Larsen (2005) and Larsen *et al.* (2010). *Intra*- and *inter*-assay CV were in all cases below five and eight percent, respectively, for both low and high control samples. The analysis of milk metabolites and enzymes was carried out at the Department of Animal Science, Aarhus University, Denmark.

Derived measures

Feed samples were collected weekly from UK, DK and IE were analysed for Net Energy for Lactation (**NE_L**) in a single run at Cumberland Valley Agricultural Services, Maryland. EBAL (in MJ/day), derived from NE_L, was determined according to the National Research Council (NRC, 2001): EBAL = NE_L feed intake - NE_L milk production - NE_L maintenance. From this, the energy input from daily dry matter intake (**DMI**) was calculated by multiplying the weekly NE_L with the observed DMI. Daily measures of milk yield were combined with the less frequent analyses of fat, protein and lactose content using the closest composition measure forwards in time to obtain the NE_L used for milk production. EBAL was only calculated if both morning and evening yield were available for the current day. Afterwards, three days (i.e. +/- 1

DIM) moving averages of EBAL were calculated and used for the analyses (Supplementary Table S1). The average BW within calendar week was used to smooth large day-to-day variation and measurement errors of scales.

PI-index was calculated as $[\log_{10}(\text{NEFA})] + [\log_{10}(\text{BHB})] - [\text{glucose}]$ (Moyes *et al.*, 2013), where plasma concentrations of the individual metabolites were standardised to an overall mean of zero and variance of one (as indicated by square brackets).

Statistical analysis

Statistical analyses and calculations were carried out using R 3.4.4 (R Core Team, 2018) and a 5% level of significance was chosen. Descriptive statistics were calculated as mean, SD, minimum, maximum, and quartiles. Multilevel random regression models, with the levels herd/diet and cows were used to estimate the variance components in the potential milk and blood biomarkers adjusted for DIM and parity. In the models for EBAL and milk metabolites and milk enzymes, DIM was included as quadratic term and all two-way interactions were included. Similar models with DIM as a 2-level factor (DIM14 and DIM35) were used to estimate the variance components in the potential blood biomarkers and PI-index.

The Intra class Correlation Coefficient (ICC) was calculated as the proportion of the total variance which could be attributed to a specific level (herd/diet or cow). For example:

$$ICC_{cow} = Var_{cow} / (Var_{herd/diet} + Var_{cow} + Residual)$$

describes the proportion of the variance that can be attributed to cows, that are adjusted for the covariates. Model control was done by assessing qq-plots and plots

of residuals vs fitted values and residuals vs DIM for each herd/diet combination and parity group.

Results

Production and health data

Summary statistics for daily milk yield for each of the herds (and by diet in the case of UK and DK) are presented in Table 2. Milk yield over the seven-week period averaged 33.3 ± 9.3 kg/day, increasing from 25.8 ± 8.1 kg/day in week 1 to 36.1 ± 9.3 kg/day in week 7. Of the 234 cows on the study, 73 had a clinical diagnosis recorded (Table 3), with mastitis, metritis, retained placenta and endometritis being the predominant health issues. Eleven of the recorded diagnoses could be associated directly with metabolic disease.

The production results and cow characteristics were within the expected range for Holstein dairy cows in early lactation (Table 2). Despite some of the diets being designed to create metabolic challenges (e.g. Ketogenic ‘High Sugar’ diet in DK), the health records demonstrated that the cows had a high metabolic health status (Table 3). The limited number of clinical diseased cows would most likely only be a minor contribution to the overall variation in the biomarkers, however there could be systematic differences in detection and treatment thresholds in the different herds that could have resulted in failure to observe some cases. This also indicates that the vast majority of the observations of EBAL, milk metabolites, milk enzymes and blood metabolites were within the range of variation that could be expected from different herds with individual diets and management and other assignable causes of variation. The observations from this study can be used to estimate variance components related to herd/diet, cows and within cows related to normal process

variation. However, the process variation will be overestimated because the assignable causes (also diseases) have not been removed.

Energy Balance, milk metabolites and milk enzymes

EBAL was calculated for the three herds (UK, DK and IE) which had the necessary feed intake and ration composition data available (n=132 cows). While there were some missing values (weeks 1 and 7 in particular), the average coverage from six to 47 DIM was 82%. Average EBAL within each herd/diet combination ranged from -29.9 MJ/day in the low concentrate diet in UK to 3.1 MJ/day in the high concentrate diet in UK (with all other diets in between: Supplementary Table S1), and a standard deviation within diet from 24.9 to 33.2 MJ/day. High concentrate and high starch diets offered in UK and DK, respectively, increased EBAL. Summary statistics of EBAL with respect to parity groups are shown in Supplementary Table S2 and summary statistics in relation to DIM for the milk metabolites and enzymes are shown in Supplementary Tables S3 and S4.

In Table 4, the variance components and ICCs from the multilevel regression models of milk metabolites, milk enzymes and the composite EBAL measure are shown, when DIM and parity are accounted for. The ICC for herd/diet ranged from 12% for Glucose-6-phosphate and LDH to 46% for urea. The ICC for cow ranged from 17% for urea to 48% for NAGase. The percentage of the total variance that can be explained by herd/diet or cow effects (sum of $ICC_{herd/diet}$ and ICC_{cow}), ranged from 42% for isocitrate to 63% for urea, with the majority around 50%. The ratio between herd/diet variance and cow variance ranged from 0.30 for LDH and NAGase to 2.55 for urea. A high ratio indicates that herd/diet is a far more important source for variation than the individual cow. Supplementary Figures S1-S9 provide additional

348 description and model control, herd/diet and parity predictions, together with
349 residuals vs. fitted values and residuals vs. DIM for overall EBAL and each of the
350 individual milk biomarkers. The figures show the overall predicted trends in the milk
351 biomarkers between DIM 1-50 and the average difference between herd/diet and
352 parity. The residual plots of residuals vs. fitted values demonstrate outliers in many of
353 the biomarkers and issues related to variance homogeneity between parity groups
354 and between different herd/diet combinations and in some plots trends and funnel
355 shapes were observed. The qq-plots (not shown) showed some issues with the
356 normality assumptions especially for BHB, LDH and NAGase, and that $\log_{10}$-
357 transformation did not entirely solve this.
358
359 *Blood metabolites, IGF-1 and Physiological Imbalance-index*
360 Descriptive statistics of the blood metabolites, IGF-1 and PI-index measured at
361 approximately 14 and 35 DIM are presented in Supplementary Table S5. The
362 variance components and the ICCs related to herd/diet and cow are given in Table 5.
363 The ICC for herd/diet ranged from 17% for cholesterol to 47% for urea, when
364 DIM and parity were accounted for. The ICC for cows ranged from 23% for urea to
365 54% for cholesterol. The percentage of the total variance that could be explained by
366 herd/diet or cow effects (sum of $ICC_{herd/diet}$ and ICC_{cow}), ranged from 56% for NEFA
367 to 75% for IGF-1, with the majority around 60%. The ratio between herd/diet variance
368 and cow variance ranged from 0.31 for cholesterol to 2.03 for urea. The models were
369 checked by boxplots of the residuals on the two time points and plots of residual vs.
370 fitted values and no obvious deviations were found.
371 Comparing the ICCs from milk vs blood for urea, glucose and BHB showed
372 that the ICCs were roughly equal, except for the herd/diet ICC for BHB (20% for milk

vs 46% for blood). Generally, the clustering of the observations (correlation between observations) within herds and cows are able to explain a larger proportion of the total variance in blood than in milk.

Discussion

In this study less than 5% of the cows developed metabolic diseases across a range of very different feeding, housing and management conditions. The disease incidence could be underestimated because disease detection followed the normal management in the herds. However, these were research herds and would generally be considered as well managed and it is not unlikely that these cows are robust and are not easily “pushed” towards physiological imbalance by diet challenges. We now assume that the diseased cows only contribute little to the variation in the biomarkers because of the fairly low observed disease incidence. We could have chosen to remove the diagnosed cows from the data/analysis but we prefer this more transparent approach of not reducing data based on uncertain diagnostic criteria.

We estimated the variance components related to herd/diet, cow and within cow/random variation adjusted for parity and DIM at sampling. These two adjustments reduced the amount of variation at a cow-level (parity adjustment) and within cows (DIM at sampling). Generally the proportion of the variation that could be associated with cow or herd/diet were higher for the blood metabolites than milk metabolites and milk enzymes, which may in part be due to the limitation of having only two observations per cow in the study period for the blood metabolites. The ICC for herd/diet for plasma urea and BHB were 0.47 and 0.46, indicating that almost half of the variation in these observations should be associated with herd/diet factors common to all cows in the herd. For plasma urea and BHB, the ratio between

herd/diet variance and cow variance were 1.6 and 2.0, and the total proportion of variance associated with either herd/diet or cow was 70% and 74% respectively. The other blood metabolites showed proportions between herd/diet variance and cow variance around and below 1, which indicates that cow variation was relatively more important. For urea the results were as expected, because dietary protein intake influences blood urea (Carroll *et al.*, 1988) and there is substantial variation in blood urea in relation to feeding (Gustafsson and Palmquist, 1993), which would provide a distribution of the variation as found here. The descriptive boxplots of $\log_{10}$(BHB) in blood (Supplementary Figure S10), highlight substantial variation between herd/diet, but also that the distribution of the observations within herd/diet were almost identical at 14 and 35 DIM. A commonly applied threshold for BHB in blood is 1.2 mM (0.08 on $\log_{10}$-scale) for subclinical ketosis in both practice and academia. There are cows that cross this threshold in DK (High Sugar diet), BE and DE. Comparing these results with the EBAL estimated from UK, DK and EI (Supplementary Table S1) suggest that it is not only EBAL that determines blood BHB concentrations, since there are no high BHB levels in the herd/diet combination with the highest negative EBAL, and none in the DK standard diet, with the latter very similar in EBAL to the 'High Sugar' diet, that were designed to be ketogenic. These results suggest that fixed thresholds for BHB in blood can be problematic to describe the metabolic status of a cow.

For metabolites and enzymes measured in milk, less of the variation could be attributed to herd/diet or cow. Urea in milk followed the same pattern as urea in blood, whereas for BHB the $ICC_{herd/diet}$ was reduced from 46% in blood to 20% in milk. For milk BHB the total amount of variation that could be attributed to herd/diet or cow was 54%. Because there are diurnal fluctuations in blood BHB

relative to feeding (Quiroz-Rocha *et al.*, 2010) and also within total mixed ration based systems (Nielsen *et al.*, 2003), it could be hypothesized that milk BHB provides a more robust indicator of the metabolic challenges in the cow. In addition, the relatively high $ICC_{herd/diet}$ for BHB in blood might be explained by variation between blood sampling time points between herds (morning vs. afternoon).

Glucose in milk is not produced by the mammary epithelia cells but is transported directly from the blood stream. However, the mammary epithelia also utilise glycose for multiple metabolic pathways, including conversion to lactose (Annison, 1983). Larsen and Moyes (2015) also found large variations in the glucose and glucose-6-phosphate in milk, but also stressed that more work is needed to identify the mechanisms that can relate these metabolites to disease risk. Residual analysis of the relations between the same blood and milk metabolites could be a useful approach to understand these relations in more details including the variation related to herd/diet.

The enzymes LDH and NAGase are enzymes that are commonly used as udder health indicators. In this study we found $ICC_{herd/diet}$ of 12% and 15% for LDH and NAGase, and the ratio between herd/diet and cow variance was 0.3 for both. These results are in line with the study of Åkerstedt *et al.* (2011), who also suggested that the use of LDH as an indicator of mastitis requires adjustment for the individual cow (and quarter). The variance that is attributed to the herd/diet was the lowest found in this study, and is likely related to differences in mastitis incidence between the different herds.

The results presented here describe biomarkers in early lactation cows in a range of different environments, management and feeding strategies. The variance components can then be considered an estimate of normal variation in the

448 biomarkers, when parity and DIM are accounted for. However, as demonstrated by
449 the supplementary figures S1-S9 accounting for herd/diet, parity and DIM did not
450 explain all the variation in the data, which indicates that other (unobserved) sources
451 of variations contribute to the residual variation. The number of outliers, the
452 differences in residual variance between parity groups and diets and trends in the
453 residuals suggest that this is not only true random error of the production process.
454 We have shown that herd and DIM contribute to the total variation in the biomarkers,
455 but the residual plots indicate that other sources of variation exist as well. A potential
456 source of variation could be physiological imbalance and these deviations could be
457 considered predictors of disturbed a production process. If these biomarkers should
458 be used for monitoring at an individual cow level, these additional sources of
459 variation should be identified and removed, thereby reducing total variation and
460 improving the predictability of the biomarker. Other diseases could be potential
461 assignable causes of variation in the biomarkers. Identification of sources of variation
462 will be a continuous process involving carefully scrutinizing outliers and deviations in
463 variation in each of the biomarkers. Based on this study, NAGase and LDH are not
464 heavily influenced by herd factors so these could be considered useful biomarker for
465 mastitis on cow level and useful for PLF. The other biomarkers, like BHB, is
466 influenced by herd factors which suggests that thresholds should be set locally to
467 describe the metabolic status of the cows and further work is needed to do this.
468 However in general, the larger the difference in a biomarker between healthy and the
469 metabolically challenged populations, the greater the amount of variation that can be
470 accepted within the healthy population. The variance components as depicted here
471 can be used to provide some guidelines into what should be considered exceptional
472 variation. Variation associated with the cow and herd/diet level are of importance,

because these cannot just be accounted for by additional recordings of the individual. In addition, a common threshold for intervention cannot be applied unless there is a large difference between the healthy and imbalanced populations. Variation in the biomarkers between the cows are most likely the most problematic source of variation, because this necessitates multiple observations for each cow. Between-cow variation is described for acetoacetate and BHB as 'individual sensitivity for hyperketonemia' in the appearance of clinical signs of ketosis (Andersson, 1984). Herd/diet variations would be less of a challenge, because the PLF tool would only need to be adjusted for each herd and/or diet. In addition, the variances in the biomarkers presented here could be beneficial in simulation studies to estimate potential impact of PLF with varying prevalences of metabolic disease.

Conclusions

The ICCs for the milk metabolites and blood metabolites demonstrated that varying proportions of the variance could be associated with herd/diet and individual cow. ICC related to the herd-diet combination were generally higher for blood metabolites than milk metabolites. These results provide valuable information about the variation in a range of biomarkers, which can be used for subsequent simulations to assess the potential of the biomarkers in PLF.

Author's contribution

MAC, NG, GO, FD, DCW and KLI designed the overall study and conceptual design. The animal experiments were supervised/or conducted by MH, MS, MTS, CM, FS, FN, FB, FC, AV, MAC and CPF in the different countries. The laboratory analyses were done by CG, TL and EM. All statistical analyses were done by LF. Earlier

versions of this manuscript were drafted by LF, MH, MS, CPF, DCW and CG including comprehensive data control. MAK and LF drafted a major revision and did additional statistical analyses on the data. All authors discussed the results and commented on the final manuscript.

Acknowledgements

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 613689. Additional information about the GenotypePlusEnvironment project can be found at http://GplusE.eu. The views expressed in this publication are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

Additional funding was provided by the respective national research agencies and the Belgian authors thank the European Union's INTERREG NWE program and the Ministry of Agriculture of Walloon Region of Belgium (Service Public de Wallonie Agriculture) for their financial support through the OptiMIR and the HappyMoo projects.

Barn staff, lab technicians and data managers are acknowledged for their invaluable work. Dr. L.J. Spicer, Oklahoma State University, is acknowledged for assistance with the IGF-1 radioimmunoassay and Dr. Parlow, the National Hormone & Peptide Program (NHPP), for supplying the anti-IGF-I, NHPP-NIDDK.

Declaration of interest

There is no direct financial interest of the authors and affiliations in the subject matter discussed in the manuscript. All financial support is identified in the Acknowledgements.

Ethics statement

The experiments were carried out in accordance with the standards recommended by the EU Directive 2010/63/EU for animal experiments.

Software and data repository resources

None of the data were deposited in an official repository.

References

Andersson L 1984. Concentrations of Blood and Milk Ketone Bodies, Blood Isopropanol and Plasma Glucose in Dairy Cows in Relation to the Degree of Hyperketonaemia and Clinical Signs. Zentralblatt für Veterinärmedizin Reihe A, 31, 683-693.

Annison EF 1983. Metabolite utilization by the ruminant mammary gland. In Biochemistry of lactation (ed. TB Mepham), pp. 399–436. Elsevier, Amsterdam, Netherlands.

Beltman ME, Forde N, Furney P, Carter F, Roche JF, Lonergan P and Crowe MA 2010. Characterisation of endometrial gene expression and metabolic parameters in beef heifers yielding viable or non-viable embryos on day 7 after insemination. Reproduction, Fertility, and Development 22, 987-999.

Berckmans D 2017. General introduction to precision livestock farming. Animal Frontiers 7, 6-11.

Berry DP, Horan B, O'Donovan M, Buckley F, Kennedy E, McEvoy M and Dillon P 2007. Genetics of grass dry matter intake, energy balance, and digestibility in grazing Irish dairy cows. Journal of Dairy Science 90, 4835-4845.

Bjerre-Harpoth V, Storm AC, Vestergaard M, Larsen M and Larsen T 2016. Effect of postpartum propylene glycol allocation to over-conditioned Holstein cows on concentrations of milk metabolites. The Journal of Dairy Research 83, 156-164.

Carroll DJ, Barton BA, Anderson GW, Smith RD 1988. Influence of Protein Intake and Feeding Strategy on Reproductive Performance of Dairy Cows. Journal of Dairy Science 71, 3470-3481.

Chagas LM, Bass JJ, Blache D, Burke CR, Kay JK, Lindsay DR, Lucy MC, Martin GB, Meier S, Rhodes FM, Roche JR, Thatcher WW and Webb R 2007. Invited review: New perspectives on the roles of nutrition and metabolic priorities in the subfertility of high-producing dairy cows. Journal of Dairy Science 90, 4022-4032.

Grelet C, Fernández Pierna JA, Dardenne P, Soyeurt H, Vanlierde A, Colinet F, Gengler N, Baeten V and Dehareng F 2016. Development of Fourier transform mid-infrared

560 calibrations to predict acetone, β-hydroxybutyrate and citrate contents in bovine milk
561 through a European dairy network. Journal of Dairy Science 99, 4816-4825.
562 Gustafsson AH and Palmquist DL 1993. Diurnal variation of rumen ammonia, serum urea,
563 and milk urea in dairy cows at high and low yields. Journal of Dairy Science 76, 475-
564 484.
565 Huzzey JM, Nydam DV, Grant RJ and Overton TR 2012. The effects of overstocking Holstein
566 dairy cattle during the dry period on cortisol secretion and energy metabolism. Journal
567 Dairy Science 95, 4421-4433.
568 Ingvartsen KL 2006. Feeding- and management-related diseases in the transition cow:
569 physiological adaptations around calving and strategies to reduce feeding-related
570 diseases. Animal Feed Science and Technology 126, 175-213.
571 Ingvartsen KL and Moyes K 2013. Nutrition, immune function and health of dairy cattle.
572 Animal 7, 112–122.
573 Jensen AL, Petersen MB and Houe, H 1993. Determination of the fructosamine
574 concentration in bovine serum samples. Journal of Veterinary Medicine Series A 40,
575 111-117.
576 Larsen T 2005. Determination of lactate dehydrogenase (LDH) activity in milk by a
577 fluorometric assay. The Journal of Dairy Research 72, 209-216.
578 Larsen T 2014. Fluorometric determination of free and total isocitrate in bovine milk. Journal
579 of Dairy Science 97, 7498-7504.
580 Larsen T 2015. Fluorometric determination of free glucose and glucose 6-phosphate in cows'
581 milk and other opaque matrices. Food Chemistry 166, 283-286.
582 Larsen T and Moyes KM 2010. Fluorometric determination of uric acid in bovine milk. The
583 Journal of Dairy Research 77, 438-444.
584 Larsen T and Moyes KM 2015. Are free glucose and glucose-6-phosphate in milk indicators
585 of specific physiological states in the cow? Animal 9, 86-93.
586 Larsen T and Nielsen NI 2005. Fluorometric determination of b-hydroxybutyrate in milk and
587 blood plasma. Journal of Dairy Science 88, 2004-2009.

588 Larsen T, Rontved CM, Ingvartsen KL, Vels L and Bjerring M 2010. Enzyme activity and
589 acute phase proteins in milk utilized as indicators of acute clinical E. coli LPS-induced
590 mastitis. Animal 4, 1672-1679.
591 Mair B, Drillich M, Klein-Jöbstl D, Kanz P, Borchardt S, Meyer L, Schwendenwein I and
592 Iwersen M. 2016Glucose concentration in capillary blood of dairy cows obtained by a
593 minimally invasive lancet technique and determined with three different hand-held
594 devices. BMC Veterinary Research 12:34. doi: 10.1186/s12917-016-0662-3.
595 Mertens K, Decuypere E, De Baerdemaeker J and De Ketelaere B 2011. Statistical control
596 charts as a support tool for the management of livestock production. The Journal of
597 Agricultural Science, 149(3), 369-384.
598 Moyes KM, Bendixen E, Codrea MC and Ingvartsen KL 2013. Identification of hepatic
599 biomarkers for physiological imbalance of dairy cows in early and mid lactation using
600 proteomic technology. Journal of Dairy Science 96, 3599-3610.
601 National Research Council (NRC) 2001. Nutrient requirements of dairy cattle, volume 1, 7th
602 edition. National Academies Press, Washington, DC, USA.
603 Nielsen NI, Ingvartsen KL and Larsen T 2003. Diurnal variation and the effect of feed
604 restriction on plasma and milk metabolites in TMR-fed dairy cows. Journal of Veterinary
605 Medicine Series A 50, 88-97.
606 Nielsen NI, Larsen T, Bjerring M and Ingvartsen KL 2005. Quarter Health, Milking Interval,
607 and Sampling Time During Milking Affect the Concentration of Milk Constituents.
608 Journal of Dairy Science 88, 3186-3200.
609 O'Leary-Kelly SW and Vokurka, RJ 1998. The empirical assessment of construct validity.
610 Journal of Operations Management 16, 387-405
611 Ospina PA, Nydam DV, Stokol T and Overton TR 2010. Evaluation of nonesterified fatty
612 acids and β-hydroxybutyrate in transition dairy cattle in the northeastern United States:
613 Critical thresholds for prediction of clinical diseases. Journal of Dairy Science 93, 546-
614 554.

615 Quiroz-Rocha GF, LeBlanc SJ, Duffield TF, Jefferson B, Wood D, Leslie KE and Jacobs RM
616 2010. Short communication: Effect of sampling time relative to the first daily feeding on
617 interpretation of serum fatty acid and β-hydroxybutyrate concentrations in dairy cattle.
618 Journal of Dairy Science 93, 2030-2033.
619 R Core Team 2018. R: A language and environment for statistical computing. R Foundation
620 for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
621 Seifi HA, LeBlanc SJ, Leslie KE and Duffield T 2011. Metabolic predictors of post-partum
622 disease and culling risk in dairy cattle. The Veterinary Journal 188, 216-220.
623 Stengärde L, Holtenius K, Tråvén M, Hultgren J, Niskanen R and Emanuelson U 2010. Blood
624 profiles in dairy cows with displaced abomasum. Journal of dairy science 93, 4691-
625 4699.
626 Wathes DC, Cheng Z, Bourne N, Taylor VJ, Coffey MP and Brotherstone S 2007.
627 Differences between primiparous and multiparous dairy cows in the inter-relationships
628 between metabolic traits, milk yield and body condition score in the periparturient
629 period. Domestic Animal Endocrinology 33, 203-225.
630 Åkerstedt M, Forsbäck L, Larsen T and Svennersten-Sjaunja K 2011. Natural variation in
631 biomarkers indicating mastitis in healthy cows. Journal of Dairy Research 78, 88-96.
632

633 **Table 1** *Overview of the diets fed to the 234 Holstein dairy cows within the six*
634 *experimental herds*

Herd	Ingredients	Diet	N cows
UK	Three iso-nitrogenous diets comprising mixtures of grass silage and concentrate in different ratios on a dry matter basis.	Low C: 30% Standard C: 50% High C: 70%	20 20 21
DK	Three iso-nitrogenous and iso-calorific diets comprising grass silage, maize silage, sugar beet pulp pellets, and concentrate including high level of barley in the "High starch" diet and high level of dextrose in the "High sugar" diet. The 'High starch' diet was intended to induce acidosis and the 'High Sugar' diet was intended to induce ketosis.	High starch: 54% C High sugar: 54% C Standard: 49% C	11 10 14
IE	A standard diet comprising grass silage, maize silage, sugar beet pulp pellets, and concentrate. In addition, each cow was offered 8 kg of concentrate per day in the parlour at milking.	Standard: 20% C	36
BE	A standard diet changing over time to include summer grazing. The standard diet comprised grass silage, maize silage, and concentrate. Moreover, cows were offered 1 kg concentrate per 2.5 l milk above 25 l/day with a maximum of 6 kg C/day at milking.	Standard: 17% C	31
DE	A standard diet comprising grass silage, maize silage, and concentrate.	Standard: 50% C	26
IT	A standard diet comprising sorghum silage, alfalfa hay, meadow hay, and concentrate.	Standard: 30% C	45

635 C: concentrate

636

637 **Table 2** *Summary statistics for daily milk yield (kg/day) of Holstein dairy cows over*
638 *the study period (1 to 50 days-in-milk (DIM))*

Herd	Diet	Mean (SD)	Median (Q1 ; Q3)	Min ; Max	N_{cows}	$N_{samples}$
UK	Low C	28.5 (7.4)	27.9 (22.8 ; 33.4)	14.0 ; 56.1	20	820
	Standard C	32.6 (9.3)	33.1 (25.1 ; 39.6)	7.9 ; 57.0	20	859
	High C	37.2 (10.3)	38.9 (29.2 ; 45.0)	7.7 ; 64.5	21	870
	Pooled	32.8 (9.8)	32.6 (25.4 ; 40.2)	7.7 ; 64.5	61	2549
DK	High starch	37.7 (10.7)	36.8 (28.9 ; 44.6)	13.0 ; 63.1	11	465
	High sugar	35.4 (7.1)	36.4 (30.6 ; 40.3)	14.0 ; 51.3	10	452
	Standard	39.7 (9.4)	39.5 (33.6 ; 47.2)	7.5 ; 61.8	14	609
	Pooled	37.8 (9.4)	37.6 (31.5 ; 44.3)	7.5 ; 63.1	35	1526
IE	Standard	33.0 (7.1)	33.3 (29.1 ; 37.9)	10.4 ; 52.4	36	1442
BE	Standard	30.5 (8.6)	30.1 (24.6 ; 35.9)	7.0 ; 62.2	31	1324
DE	Standard	38.6 (8.5)	39.6 (34.9 ; 44.4)	2.7 ; 62.3	25	1139
IT	Standard	29.8 (8.0)	30.3 (24.1 ; 35.8)	3.3 ; 50.8	43	2089
All	Pooled	33.3 (9.3)	33.6 (26.8 ; 39.5)	2.7 ; 64.5	231	10 069

639 C: concentrate, Q1: first quartile, Q3: third quartile.

640

Table 3 *Clinical diagnoses of Holstein dairy cows during the period from calving to 50 Days in Milk (73 cows had a clinical diagnosis, while 161 cows had no clinical diagnoses recorded)*

ICAR Term [1]	ICAR code	Occurrence
Anoestria	2.05.02.01.02.	4
Bronchopneumonia	1.06.07.06.	8
Digital dermatitis	1.10.07.10.	1
Displaced abomasum	1.07.12.05.	6
Endometritis	2.05.01.01.	19
Interdigital hyperplasia	1.10.06.10.	4
Lameness	1.09.05.	9
Mastitis	1.13.	43
Metabolic diseases and deficiencies	6.	1
Metritis	2.04.05.02.	20
Milk fever	6.03.01.01.	4
Peritonitis	1.07.14.03.	3
Retained placenta	2.04.03.	20
Sole ulcer	1.10.07.03.	1
Total (n=73 cows)		143

[1] ICAR terms sorted alphabetically.

646

647 **Table 4** *Variance components and Intra Class Correlations (ICC) on Energy Balance*

648 *(EBAL) and milk metabolites and enzymes from early lactation Holstein dairy cows.*

649 In the model variation that could be attributed to parity and lactation stage is

650 removed.

Model	Herd/diet variance	Cow variance	Residual	$ICC_{herd/diet}$	ICC_{cow}	Herd/diet variance : Cow variance
EBAL[1]	125	289	389	0.15	0.36	0.43
Uric Acid	611	687	1746	0.20	0.23	0.89
Urea	0.84	0.33	0.66	0.46	0.17	2.55
Isocitrate	0.0003	0.0005	0.0011	0.18	0.24	0.60
log_{10}(BHB[2])	0.0096	0.0163	0.0221	0.20	0.34	0.59
Glucose	0.0024	0.0026	0.0033	0.29	0.31	0.92
Glucose-6-phospate	0.0005	0.0015	0.0020	0.12	0.37	0.33
log_{10}(LDH[3])	0.0078	0.0263	0.0325	0.12	0.39	0.30
log_{10}(NAGase4)	0.0066	0.0220	0.0169	0.15	0.48	0.30

651 [1]: Energy Balance was calculated for three herds (UK, DK and IE) and 7 diets.
652 [2]: β-hydroxybutyrate.
653 [3]: *N*-acetyl-β-D-glucosaminidase.
654 [4]: lactate dehydrogenase.
655
656

Table 5 *Variance components and Intra Class Correlations (ICC) on Physiological Imbalance-index (PI-Index), IGF-1 and blood metabolites from early lactation Holstein dairy cows.* In the model variation that could be attributed to parity and lactation stage is removed.

Model	Herd/diet variance	Cow variance	Residual	$ICC_{herd/diet}$	ICC_{cow}	Herd/diet variance : Cow variance
PI-index	1.59	1.65	1.44	0.34	0.35	0.96
IGF-1	0.032	0.030	0.021	0.39	0.36	1.07
Urea	0.65	0.32	0.42	0.47	0.23	2.03
log_{10}(BHB[1])	0.026	0.016	0.014	0.46	0.28	1.63
log_{10}(NEFA[2])	0.019	0.032	0.042	0.21	0.35	0.59
Glucose	0.064	0.068	0.086	0.29	0.31	0.94
Fructosamine	94.4	101.4	134.0	0.29	0.31	0.93
Cholesterol	0.13	0.42	0.23	0.17	0.54	0.31

[1]: β-hydroxybutyrate.
[2]: non-esterified fatty acid.